The Hidden Fortress: Structure and substructure of the complex strong lensing cluster SDSS J1029+2623

Masamune Oguri,[1][*] Tim Schrabback,[2,3] Eric Jullo,[4] Naomi Ota,[5]
Christopher S. Kochanek,[6] Xinyu Dai,[7] Eran O. Ofek,[8] Gordon T. Richards,[9]
Roger D. Blandford,[3] Emilio E. Falco[10] and Janine Fohlmeister[11]

[1] *Kavli Institute for the Physics and Mathematics of the Universe (Kavli IPMU, WPI), University of Tokyo, Chiba 277-8583, Japan*
[2] *Argelander-Institut für Astronomie, Auf dem Hügel 71, 53121 Bonn, Germany*
[3] *Kavli Institute for Particle Astrophysics and Cosmology, Stanford University, 382 Via Pueblo Mall, Stanford, CA 94305-4060, USA*
[4] *Aix Marseille Université, CNRS, LAM (Laboratoire d'Astrophysique de Marseille) UMR 7326, 13388, Marseille, France*
[5] *Department of Physics, Nara Women's University, Kitauoyanishimachi, Nara, Nara 630-8506, Japan*
[6] *Department of Astronomy, The Ohio State University, Columbus, OH 43210, USA*
[7] *Homer L. Dodge Department of Physics and Astronomy, University of Oklahoma, Norman, OK 73019, USA*
[8] *Benoziyo Center for Astrophysics, Weizmann Institute of Science, 76100 Rehovot, Israel*
[9] *Department of Physics, Drexel University, 3141 Chestnut Street, Philadelphia, PA 19104, USA*
[10] *Harvard-Smithsonian Center for Astrophysics, Cambridge, MA02138, USA*
[11] *Astronomisches Rechen-Institut, Zentrum für Astronomie der Universität Heidelberg, Mönchofstr. 12-14, 69120 Heidelberg, Germany*

ABSTRACT

We present *Hubble Space Telescope (HST)* Advanced Camera for Surveys (ACS) and Wide Field Camera 3 (WFC3) observations of SDSS J1029+2623, a three-image quasar lens system produced by a foreground cluster at $z = 0.584$. Our strong lensing analysis reveals 6 additional multiply imaged galaxies in addition to the multiply imaged quasar. We confirm the complex nature of the mass distribution of the lensing cluster, with a bimodal dark matter distribution which deviates from the *Chandra* X-ray surface brightness distribution. The Einstein radius of the lensing cluster is estimated to be $\theta_{\rm E} = 15''2 \pm 0''5$ for the quasar redshift of $z = 2.197$. We derive a radial mass distribution from the combination of strong lensing, *HST*/ACS weak lensing, and Subaru/Suprime-cam weak lensing analysis results, finding a best-fit virial mass of $M_{\rm vir} = 1.55^{+0.40}_{-0.35} \times 10^{14} h^{-1} M_{\odot}$ and a concentration parameter of $c_{\rm vir} = 25.7^{+14.1}_{-7.5}$. The lensing mass estimate at the outer radius is smaller than the X-ray mass estimate by a factor of ~ 2. We ascribe this large mass discrepancy to shock heating of the intracluster gas during a merger, which is also suggested by the complex mass and gas distributions and the high value of the concentration parameter. In the *HST* image, we also identify a probable galaxy, GX, in the vicinity of the faintest quasar image C. In strong lens models, the inclusion of GX explains the anomalous flux ratios between the quasar images. The morphology of the highly elongated quasar host galaxy is also well reproduced. The best-fit model suggests large total magnifications of 30 for the quasar and 35 for the quasar host galaxy, and has an AB time delay consistent with the measured value.

Key words: dark matter — galaxies: clusters: individual: SDSS J1029+2623 — gravitational lensing: strong — gravitational lensing: weak — quasars: individual: SDSS J1029+2623 — X-rays: galaxies: clusters

1 INTRODUCTION

SDSS J1029+2623 (Inada et al. 2006; Oguri et al. 2008) is a unique quasar lens system with three images of the

$z_s = 2.197$ quasar produced by a foreground cluster of galaxies at $z_l = 0.584$. Such "naked cusp" configurations are predicted to comprise a significant fraction of cluster-scale quasar lenses (Oguri & Keeton 2004; Li et al. 2007; Minor & Kaplinghat 2008). All three quasar images are also detected at radio (Kratzer et al. 2011) and X-ray wave-

[*] E-mail: masamune.oguri@ipmu.jp

lengths (Ota et al. 2012), and show mini-broad absorption line (mini-BAL) features in the Lyα, SiIV, and CIV emission lines.

These earlier observations already implied that the lensing cluster is likely complex. The optical images clearly show two bright central galaxies in the lensing cluster. Follow-up spectroscopy found that the two galaxies have a velocity difference of $\Delta v \simeq 2800 \ \mathrm{km\,s^{-1}}$ (Oguri et al. 2008), which might imply a high-speed collision of two massive haloes. *Chandra* X-ray observations show that the X-ray surface mass distribution is centred on one of the central galaxies, with an additional X-ray subpeak $\sim 9''$ North-West of the main peak that has no optical counterpart (Ota et al. 2012). This complex X-ray morphology also suggests that the lensing cluster is in a merger. It is crucial to construct a reliable mass model from a lensing analysis in order to draw any conclusion on the dynamical state of the cluster.

Another interesting feature of this system is the anomalous flux ratios seen in the quasar images. Simple mass models predict the relative brightnesses of the three quasar images A, B, and C should be $B \approx C > A$, whereas optical, radio, and X-ray observations all found relative brightnesses of $B \approx A > C$. The fact that this "anomaly" is seen in radio (Kratzer et al. 2011) and absorption-corrected X-ray (Ota et al. 2012) flux ratios indicates that it is not due to extinction by any intervening matter. In addition, the rough agreement of the flux ratios between three different wavelengths implies that it is not entirely due to stellar microlensing (but see also Motta et al. 2012), particularly since microlensing generally leads to large differences between optical and X-ray flux ratios (e.g., Dai et al. 2010a). Thus the most plausible explanation for the anomalous flux ratio is millilensing by (dark halo) substructures (Kratzer et al. 2011). An accurate macro lens model is crucial for breaking the degeneracy between the macro model uncertainties and the amount of substructure required to account for the flux ratios.

In this paper, we present a new lensing analysis of SDSS J1029+2623 utilizing deep multiband *Hubble Space Telescope (HST)* observations of this system. Several multiply imaged galaxies, which are newly identified from the *HST* images, are used to construct an accurate mass model of the core of the lensing cluster. Strong lensing results are then combined with weak lensing results from the *HST* images and our earlier wide-field weak lensing analysis based on the Subaru/Suprime-cam images (Oguri et al. 2012). We then use the mass model to interpret the *Chandra* X-ray surface brightness profiles (Ota et al. 2012) and the anomalous flux ratios between the quasar images.

The structure of this paper is as follows. After briefly describing the observations in Section 2, we present the strong lensing analysis in Section 3 and the weak lensing analysis in Section 4. We then combine the various observations in Section 5 to discuss the properties of the lensing cluster. Section 6 is devoted to discussions of the anomalous flux ratio between the quasar images. Finally we summarize the results on Section 7. Throughout this paper we assume a flat cosmological model with matter density $\Omega_M = 0.275$, cosmological constant $\Omega_\Lambda = 0.725$, and dimensionless Hubble constant $h = 0.702$ (Komatsu et al. 2011). All photometry is in the AB system without correcting for Galactic extinction.
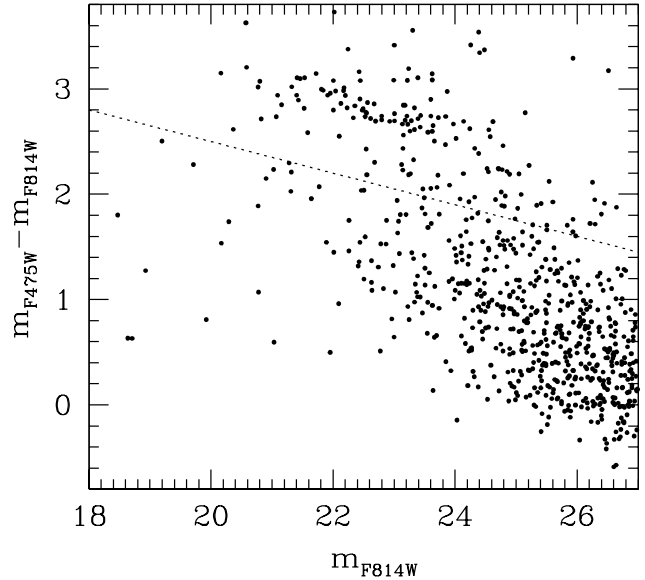

Figure 1. The colour-magnitude diagram from the *HST* ACS images. Here the magnitudes are $1''$ diameter aperture magnitudes. The dotted line indicates the colour cut used to eliminate cluster member galaxies for weak lensing analysis presented in Section 4.

2 OBSERVATIONS

We observed SDSS J1029+2623 with the Advanced Camera for Surveys (ACS) and the Wide Field Camera 3 (WFC3) on *HST* (GO-12195; PI Oguri). The observation consisted of 2 orbits (total exposure 5276 sec) of ACS/F475W imaging on 2011 May 15, 3 orbits (total exposure 8010 sec) of ACS/F814W imaging on 2011 April 17, and 2 orbits (total exposure 5223.5 sec) of WFC3/F160W imaging on 2011 April 15. We mostly follow the data reduction procedure described in Schrabback et al. (2010), which uses MultiDrizzle (Koekemoer et al. 2002) for cosmic ray removal, distortion removal and stacking. A pixel scale of $0.''05$ is adopted for both the ACS and WFC3 images. The ACS (WFC3) frames are drizzled using the lanczos3 (square) kernel, with a `pixfrac` of 1 (0.8). The ACS data are preprocessed using the pipeline CALACS, which also removes bias striping and provides a pixel-level charge transfer inefficiency correction.

We use SExtractor (Bertin & Arnouts 1996) to construct a photometric catalogue of the galaxies in the *HST* images. For each galaxy, we derive a total magnitude from the `MAG_AUTO` parameter and a $1''$ diameter aperture magnitude. We show the colour-magnitude diagram of the SDSS J1029+2623 field from the two ACS images in Figure 1. The Figure shows a clear red-sequence of passively evolving galaxies in the lensing cluster at $m_{\mathrm{F475W}} - m_{\mathrm{F814W}} \sim 3$.

We also use wide-field Subaru Suprime-cam (Miyazaki et al. 2002) images of SDSS J1029+2623 for the analysis. These images consist of 1200 sec g-band, 2700 sec r-band, and 1920 sec i-band images, with a seeing FWHM of $0.''65$ in r-band. The weak lensing analysis of the Subaru data was presented in Oguri et al. (2012), which we adopt in this paper.

SDSS J1029+2623 was also observed with *Chandra* X-

Table 1. List of multiple images for strong lensing analysis.

ID	R.A. (J2000)[a]	Decl. (J2000)[a]	redshift[b]
A	157.308139	26.388441	2.197
B	157.309398	26.394600	
C	157.309612	26.394125	
1.1	157.298104	26.390875	[1.93–2.36]
1.2	157.297918	26.392583	
1.3	157.300902	26.397536	
2.1	157.298228	26.391494	[1.84–2.27]
2.2	157.298197	26.391783	
2.3	157.301516	26.397847	
3.1	157.299103	26.392536	[2.62–3.24]
3.2	157.307476	26.391480	
3.3	157.304173	26.398397	
3.4	157.301584	26.388161	
3.5	157.301770	26.388063	
3.6	157.301878	26.388022	
3.7	157.303243	26.392097	
4.1	157.299258	26.392661	[2.58–3.26]
4.2	157.307677	26.391452	
4.3	157.304390	26.398286	
4.4	157.302359	26.387869	
4.5	157.303531	26.392152	
5.1	157.302080	26.394772	[0.89–1.02]
5.2	157.300917	26.392008	
5.3	157.300964	26.391730	
6.1	157.309073	26.395119	2.197
6.2	157.309956	26.392452	
6.3	157.308576	26.388813	

[a] See Figures 2 and 3 for the configuration of the multiple image systems and the morphology of each image.

[b] Redshift ranges in brackets are 2σ redshift ranges predicted by the strong lens models.

ray observatory (ObsID 11755; PI Oguri) on 2010 March 11 for 60 ksec. The X-ray image clearly reveals X-ray emission both from the lensing cluster and the three quasar images. The data reduction and basic spectral and image analysis were detailed in Ota et al. (2012).

3 STRONG LENSING ANALYSIS

3.1 Modelling method

The main objective of the *HST* imaging observation was to constrain the central mass distribution of the lensing cluster using strong lensing information. For this purpose, we identify multiply imaged galaxies in addition to the three quasar images based on their colours and morphologies combined with matching multiple image candidates while iteratively refining the mass models (see below for details of our mass model). In total, we identified 5 additional sets of multiply imaged galaxies (ID 1-5), which are shown in Figures 2 and 3, and summarized in Table 1. In addition, the *HST* image shown in Figure 2 exhibits a prominent host galaxy of the lensed quasar which is highly elongated due to lensing. In order to take account of the shape of the lensed host galaxy, we include an additional set of multiple images (ID 6) which roughly correspond to the edge of the host galaxy. To summarize, in this paper we use the positions of 27 multiple images of 7 systems as constraints.

We parametrically model the lens using the public software *glafic* (Oguri 2010), although in the iterative process to identify the multiple images we partly used the public software *lenstool* (Jullo et al. 2007) as well. The mass model mainly consists of dark halo components modelled by elliptical Navarro, Frenk, & White (1997, hereafter NFW) profiles with the radial profile of $\rho(r) \propto r^{-1}(r + r_s)^{-2}$ and member galaxies modelled as elliptical pseudo-Jaffe models (e.g., Cohn et al. 2001) with the radial profile of $\rho(r) \propto r^{-2}(r^2 + r_{\mathrm{cut}}^2)^{-1}$. Since the cluster core contains two bright galaxy concentrations, we include two cluster-scale dark halo components centred at galaxy G1a (R.A.=157.305789, Decl.=26.392602) and G2 (R.A.=157.302083, Decl.=26.392344). Each dark halo component has 4 parameters: the ellipticity and position angle of their isodensity contours, the virial mass M_{vir}, and the concentration parameter c_{vir}. To reduce the number of parameters, we assume that the velocity dispersion σ and the cutoff radius r_{cut} of the pseudo-Jaffe models scale with the luminosities of the galaxies as $\sigma \propto L^{1/4}$ and $r_{\mathrm{cut}} \propto L^{1/2}$, and regard the normalisations of the scaling relations as free parameters. However, we do not apply these scaling relations for one member galaxy just North of images 3.4–3.6 because the locations of these images are very sensitive to the properties of this galaxy. The velocity dispersion and cutoff radius of this particular galaxy are included as free parameters. The ellipticity and position angle of the pseudo-Jaffe model for each member galaxy are fixed to the value measured from the *HST* ACS/F814W image. In addition to the dark halo and member galaxy components, we include four perturbation terms (lens potential $\phi \propto r^2 \cos m\theta$) with $m = 2$ (external shear), 3, 4, and 5 which effectively describe the asymmetries of cluster mass distributions that are commonly seen in simulations (e.g., Meneghetti et al. 2007). Since we have no spectroscopic redshift measurements for the multiply imaged galaxies, the redshifts of the lensed galaxies are also treated as the free parameters.

We use the positions of the 27 multiple images summarized in Table 1 as observational constraints. We assume positional errors of $0\rlap{.}''1$ for the quasar images and $0\rlap{.}''4$ for the galaxy images. These errors are larger than the measurement uncertainties, and are chosen to take account of any complexity of the lens potential that is not accounted for in the parametric mass model. For the quasar images, we include flux ratios of $A/B = 0.802$ and $C/B = 0.455$ measured in the *Very Large Array* (VLA) radio image (Kratzer et al. 2011), but assuming a conservative 20% error on the flux ratios. The assumed error is larger than the measurement errors (peak flux densities of 0.206, 0.257, 0.117 mJy/beam for images A, B, and C, with an error of 0.013 mJy/beam for each component) in order to take account of time variabilities as well as the complexity of the lens potential. Recently, we measured time delay between quasar images A and B as $\Delta t_{AB} = 744 \pm 10$ days (Fohlmeister et al. 2012). We include this measurement as an observational constraint, although we assume a 5% error on Δt_{AB}, $\sigma(\Delta t_{AB}) = 37$ days, which is larger than the measurement error, because the time delay depends also on the assumed Hubble constant and is limited by cosmic variance due to matter fluctuations along the line-of-sight (e.g., Bar-Kana 1996). In total, we have 57 observational constraints and 39 model parameters (8 from the dark halo components, 4 from the member galaxy com-
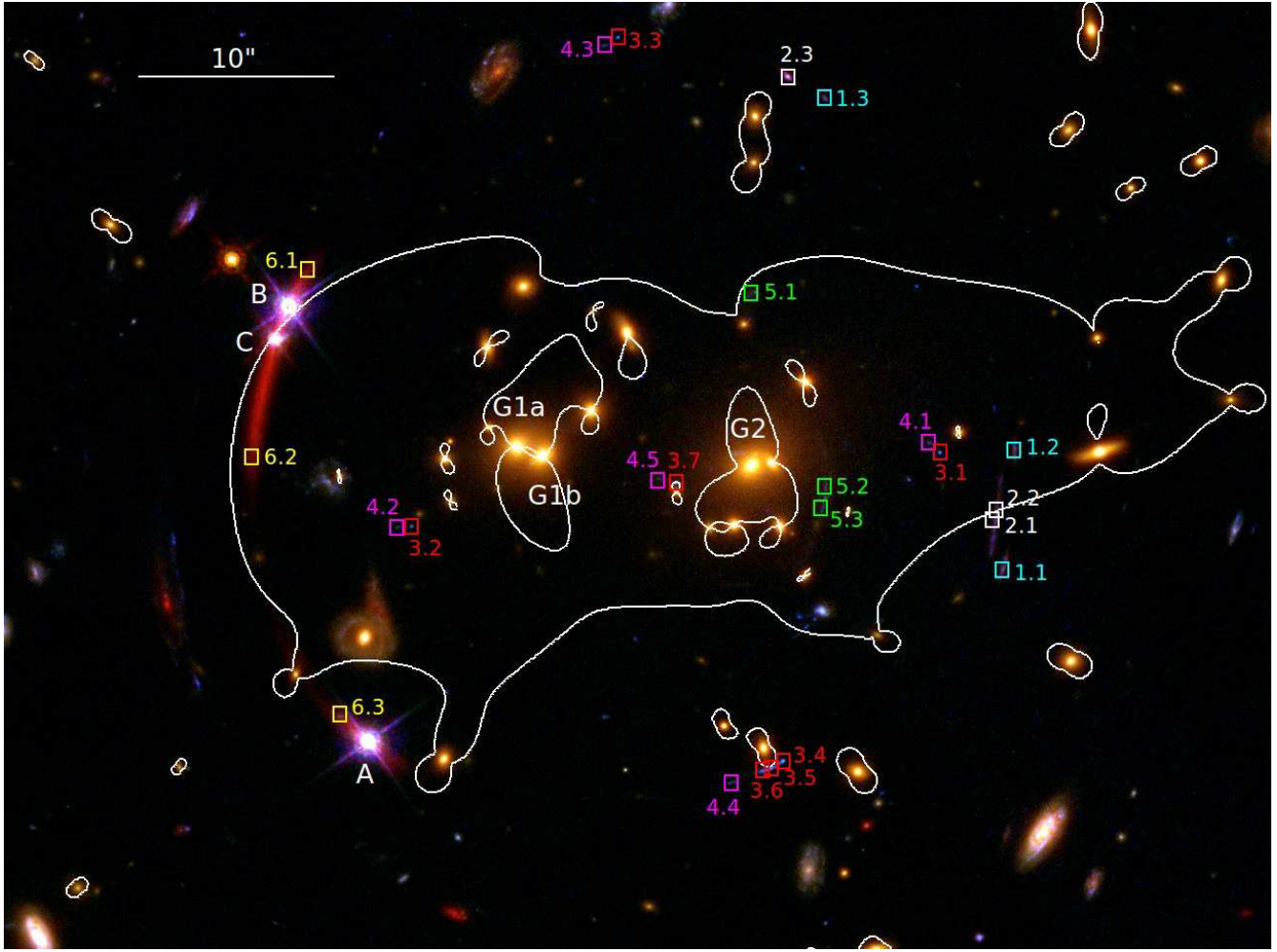
Figure 2. Colour image of the SDSS J1029+2623 field from the *HST* ACS/F475W, ACS/F814W, and WFC3/F160W observations. North is up and West is to the right. The three quasar images are labelled by A, B, and C. Squares with ID numbers indicate the locations of identified multiple images (see also Table 1). Galaxies G1a, G1b, G2 are the central galaxies of the lensing cluster. Solid lines show the critical lines predicted by our best-fit mass model at $z_s = 2.197$, the redshift of the strongly lensed quasar.
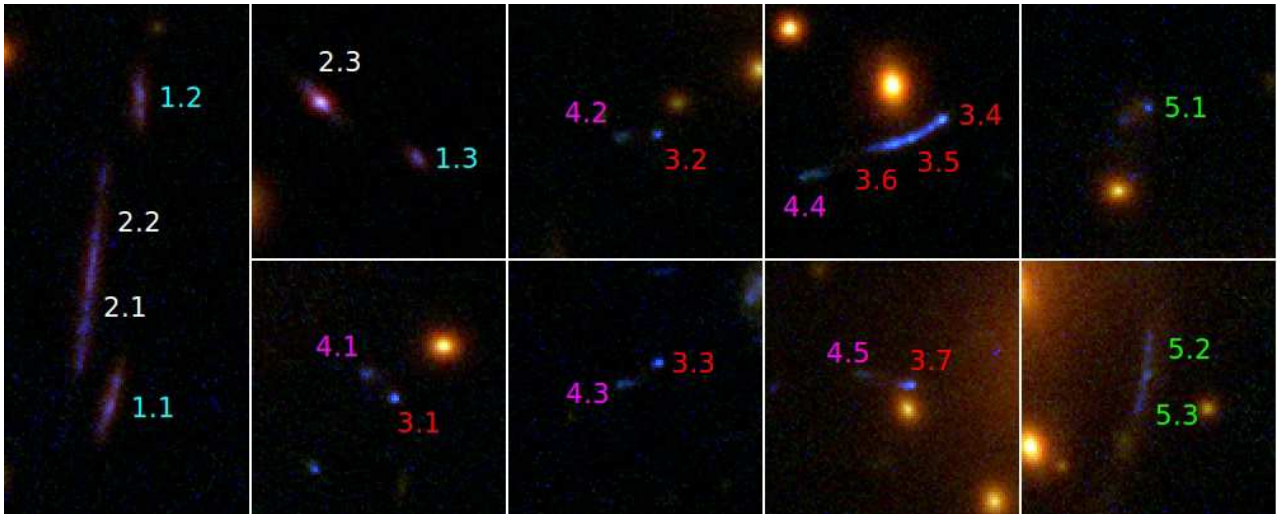
Figure 3. Similar to Figure 2, but zooming in on the multiply imaged galaxies identified in this paper.
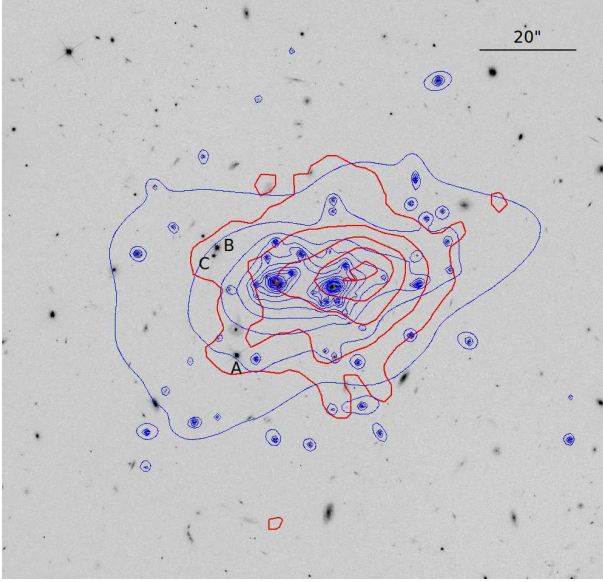
Figure 4. Contours of the best-fit surface mass density profile from strong lens modelling (*thin blue*) and the *Chandra* X-ray surface brightness profile (*thick red*) are overlaid on the *HST* ACS/F814W image. In the X-ray contours we removed the quasar images, but there are some subtraction residuals that distort the X-ray contours.

ponents, 8 external perturbations, 5 source redshifts, and 14 source positions). Thus, the models have 18 degrees of freedom.

3.2 The best-fit model

We derive the best-fit mass model by searching for the minimum χ^2. To speed-up the calculation, we estimate χ^2 in the source plane, which has been shown to be reasonably accurate for most of strong lens modelling purposes (see, e.g., Oguri 2010). We find a χ^2 for the best-fit model of $\chi^2 = 30$, which is significantly larger than the number of degrees of freedom $N_{\rm dof} = 18$. In fact, a large fraction of the χ^2 comes from the constraints on the flux ratio. While our input flux ratios are $A/B = 0.802$ and $C/B = 0.455$, the best-fit model predicts $A/B = 0.082$ and $C/B = 1.14$. This results in a large χ^2 contribution despite our conservative errors on the flux ratios, and indicates that the flux ratio anomaly found in previous simple mass models (Oguri et al. 2008; Kratzer et al. 2011) persists even in this much more complex mass model. Once the flux ratios are removed from the observational constraints, the best-fit model yields $\chi^2 = 10$ for $N_{\rm dof} = 18 - 2 = 16$ degrees of freedom, suggesting that our choice of the positional errors on the image locations is reasonable. In particular, our best-fit model has a time delay of $\Delta t_{AB} = 737$ days that reproduces the observed AB time delay almost perfectly. We explore the flux ratio anomaly further in Section 6.

The critical curves of the best-fit model are displayed in Figure 2. We find that contributions from both dark halo components to the critical curves are significant, and as a consequence, the centroid of the main critical curve responsible for the three quasar images falls between G1a and G2.

In the best-fit model, the dark halo at G2 has a smaller virial mass but a larger concentration than the dark halo at G1a (see Table 2). As a result, the halo of G2 contributes more mass inside the Einstein radius than the halo at G1a. Because the mass and concentration parameters constrained from strong lens modelling are quite degenerate with each other, for a more robust comparison, we fix the concentration parameters to $c_{\rm vir} = 5$ and find best-fit masses of $M_{\rm vir} \simeq 1.5 \times 10^{14} h^{-1} M_\odot$ for the dark halo at G2 and $M_{\rm vir} \simeq 1.0 \times 10^{14} h^{-1} M_\odot$ for the dark halo at G1a. Thus, the dark halo associated with G2 is more massive, although the effect of the dark halo at G1a is very significant. We compare the surface mass density of the best-fit model with the *Chandra* brightness profile (Ota et al. 2012) in Figure 4. As pointed out in Ota et al. (2012), the X-ray surface brightness profile of the main cluster component is centred at G2, with an additional subpeak North-West of the main peak, which appears to be different from the structure of the mass distribution found in the strong lens models. On the other hand, the position angles of the best-fit mass model, $\theta_e = 106.3$ deg (measured East of North) for the dark halo at G2 and 105.7 deg for the dark halo at G1a, are in good agreement with the position angle of the X-ray surface brightness profile, $\theta_X = 110.6 \pm 5.2$ deg (Ota et al. 2012).

An important quantity to characterise strong lens systems is the Einstein radius $\theta_{\rm E}$, although the definition of the Einstein radius is not unique, for such a complex lens system. In this paper, we define the Einstein radius as the radius such that the mean enclosed surface density equals the critical surface density,

$$\bar{\kappa}(< \theta_{\rm E}) \equiv \frac{1}{\pi \theta_{\rm E}^2} \int_{|\boldsymbol{\theta}| < \theta_{\rm E}} \kappa(\boldsymbol{\theta}) d\boldsymbol{\theta} = 1, \qquad (1)$$

adopting the average of the locations of G2 and G1a (R.A.=157.303936, Decl.=26.392473) as the centre. The Einstein radius of the best-fit model is $\theta_{\rm E} = 15''\!.2$ for the quasar redshift of $z_s = 2.197$, which is consistent with the extent of the main critical curve shown in Figure 2. However, we note that the choice of the centre is somewhat arbitrary. If we move the centre for calculating the Einstein radius by $\pm 1''$ along the North-South and East-West directions, the value changes in the range $14''\!.8 < \theta_{\rm E} < 15''\!.4$.

We explore the errors on the model parameters employing a Markov Chain Monte Carlo technique. We adopt Metropolis-Hastings sampling with a multivariate-Gaussian for the proposal distribution. From this procedure we estimate the possible range of redshifts for the multiply imaged galaxies, which we summarize in Table 1. We also estimate errors on key model parameters of the best-fit strong lens mass model, which are summarized in Table 2. For comparison, we also derive photometric redshifts of the multiply imaged galaxies based on the three-band *HST* photometry using the public code EAZY (Brammer, van Dokkum, & Coppi 2008), and find that the photometric redshifts are in good agreement with the model predictions for ID 1, 2, and 5. The photometric redshifts are lower than the model predictions for ID 2 and 3, although the upper bounds on their photometric redshifts of $z_p \lesssim 2.6$ (2σ) marginally overlap with the expected redshift ranges from strong lens modelling. In addition, we find that our strong lens modelling constrains the Einstein radius to the 2σ range of $14''\!.5 < \theta_{\rm E} < 15''\!.5$. Considering the addi-

Table 2. Best-fit values of key model parameters from the strong lens models. The position angle θ_e is measured East of North. Aside from the galaxy near image 3.4-3.6, the parameters of the other member galaxies are scaled from those for G2 as $\sigma \propto L^{1/4}$ for the velocity dispersion and $r_{\rm cut} \propto L^{1/2}$ for the cutoff radius. The uncertainties are 2σ confidence limits from the Markov Chain Monte Carlo analysis.

component	$M_{\rm vir}$ $(10^{14}h^{-1}M_\odot)$	$c_{\rm vir}$	σ $({\rm km\,s^{-1}})$	$r_{\rm cut}$ (arcsec)	e	θ_e (deg)
dark halo at G1	$1.14^{+5.35}_{-2.20}$	$5.49^{+3.91}_{-3.53}$	$\cdots$	$\cdots$	$0.74^{+0.07}_{-0.15}$	$106.3^{+1.4}_{-3.8}$
dark halo at G2	$3.04^{+1.79}_{-2.76}$	$2.94^{+7.29}_{-0.78}$	$\cdots$	$\cdots$	$0.41^{+0.12}_{-0.33}$	$105.7^{+3.8}_{-3.0}$
galaxy G2	$\cdots$	$\cdots$	291^{+98}_{-64}	$13.4^{+5.7}_{-11.2}$	$(\equiv 0.252)$	$(\equiv -57.6)$
galaxy near 3.4-3.6	$\cdots$	$\cdots$	141^{+70}_{-38}	> 1.4	$(\equiv 0.313)$	$(\equiv -167.9)$

tional uncertainties on the cluster centre mentioned above, we adopt $\theta_{\rm E} = 15''2 \pm 0''5$ (1σ error) for the combined lensing analysis in Section 5.

4 WEAK LENSING ANALYSIS

4.1 HST weak lensing measurement

We carry out the weak lensing analysis using the *HST* ACS/F814W image. We closely follow Schrabback et al. (2010) for the shear measurement procedure, which is based on the object catalogue generated by SExtractor (Bertin & Arnouts 1996) and the weak lensing method outlined in Kaiser et al. (1995). In particular, we model temporal point spread function (PSF) variations based on a principal component analysis of dense stellar fields. To correct for the systematic shear underestimates of the KSB method due to noise bias, we include an empirical calibration bias as a function of the signal-to-noise ratio of each galaxy. The method was tested against *HST* ACS-like simulation data and was shown to be accurate to better than 2% over the entire magnitude range used (see Schrabback et al. 2010, for details).

A possible systematic effect in cluster weak lensing analysis is a dilution of the weak lensing signals by member galaxies (e.g., Broadhurst et al. 2005; Medezinski et al. 2007, 2010). In addition to member galaxies in the red-sequence (see Figure 1), the galaxy catalogue may contain cluster member galaxies up to ~ 0.5 mag bluer than the red sequence (Medezinski et al. 2007). Hence, in order to reduce the dilution, we apply a colour cut of

$$m_{\rm F475W} - m_{\rm F814W} < -0.15(m_{\rm F814W} - 22) + 2.2, \qquad (2)$$

for our weak lensing analysis, where the magnitudes are the $1''$ diameter aperture magnitudes shown in Figure 1. In addition, we apply a total magnitude cut of $m_{\rm F814W} < 27$ where the number counts of the galaxies in the data turns over. The final source galaxy number density of ~ 75 arcmin^{-2} after applying these selection limits is still large enough to allow high-resolution weak lensing measurements near the cluster core.

We estimate the mean depth $\langle D_{ls}/D_{os}\rangle$ of our source galaxy sample using the redshift distribution of galaxies as a function of m_{F814W} derived by Schrabback et al. (2010). Although the colour cut (Equation 2) can in principle modify the redshift distribution, we here ignore this effect because a majority of galaxies outside the red-sequence are not removed by this colour cut (see Figure 1) and therefore the
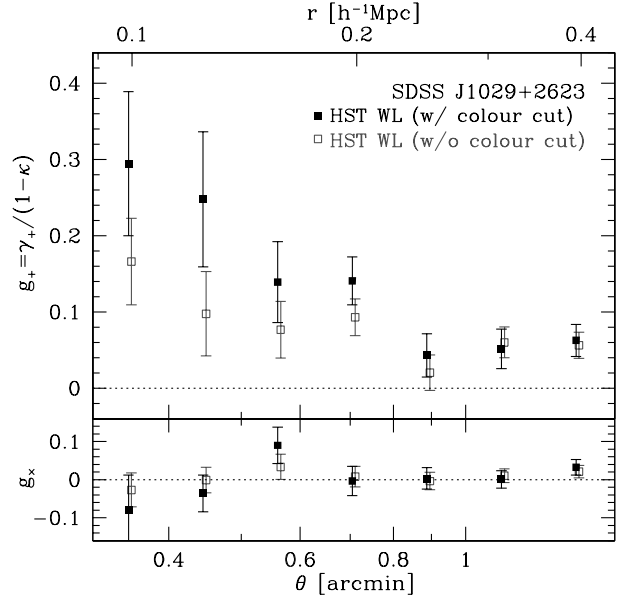

Figure 5. Tangential shear profiles from the weak lensing analysis of the *HST* ACS/F814W image. Filled squares show the shear profile after the colour cut to remove cluster member galaxies (equation 2), while open squares are the shear profile without the colour cut. The lower panel shows the radial profile of the 45° rotated component.

effect of the colour cut on the redshift distribution is expected to be small. As a cross check, we also estimate the mean depth using a photometric redshift catalogue in the Hubble Ultra Deep Field (Coe et al. 2006). After applying the same colour (Equation 2) and magnitude cuts, we obtain a mean depth that is quite consistent with the earlier estimate. We then estimate the effective source redshift at which the distance ratio becomes equal to the mean depth. We find an effective source redshift of $z_{s,{\rm eff}} = 1.06$, which we adopt for our *HST* weak lensing analysis throughout the paper.

4.2 Result

Figure 5 shows the tangential shear profile from the *HST* ACS/F814W source galaxy catalogue. Here we again adopt the average of the locations of G2 and G1a as the cluster centre, although we note that the tangential shear profile is not very different even if we adopt the location of
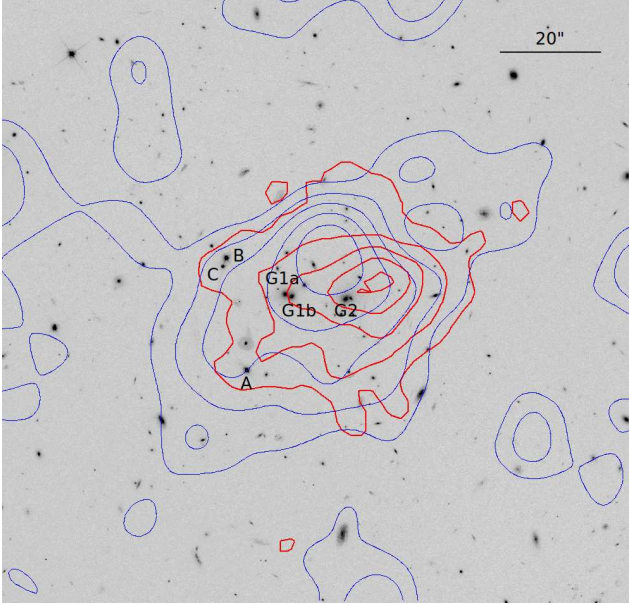
Figure 6. Similar to Figure 4, but comparing the surface mass density profile from the weak lensing analysis of the *HST* ACS/F814W image (*thin blue*) to the *Chandra* X-ray surface brightness profile (*thick red*). The mass map is Gaussian-smoothed with $\sigma = 8''$.

G1a or G2 as a cluster centre. The tangential shear detections are statistically significant, with increasing shear toward the cluster centre. To check the impact of dilution by cluster member galaxies, we derived a tangential shear profile from the source galaxy catalogue without the colour cut defined in equation (2). The resulting shear profile, which is also shown in Figure 5, has a significantly weaker amplitude compared with the profile including the colour cut, and the effect is stronger near the cluster centre. This confirms earlier claims that proper selection of galaxies to remove cluster members is essential for accurate weak lensing analysis. We also check the radial number density profile of the source galaxies and find that the number counts after the colour cut decreases toward the cluster centre, as expected from the lensing magnification effect (e.g., Umetsu & Broadhurst 2008) and masking by cluster member galaxies.

Next we explore the two-dimensional mass map reconstructed from weak lensing, using the inversion technique of Kaiser & Squires (1993). As shown in Figure 6, there is a significant weak lensing mass peak around galaxy G1a and G2, which is consistent with the strong lens modelling result shown in Section 3. For more quantitative comparisons with the *Chandra* X-ray surface brightness profile, we fit the binned, unsmoothed two-dimensional shear map with an elliptical NFW profile, following the methodology developed in Oguri et al. (2010) and Oguri et al. (2012). The model involves 6 free parameters: the virial mass and concentration parameter, the ellipticity and position angle, and the centroid. We find that the *HST* weak lensing data constrain the centroid of the surface mass density profile to $(\Delta x, \Delta y) = (1.2^{+6.2}_{-7.4}, 1.7^{+6.7}_{-8.9})$, in arcseconds from the average of the locations of G1a and G2, where positive directions of Δx and Δy correspond to West and North, respectively. Note that $(\Delta x, \Delta y) = (\mp 5.975, \pm 0.465)$ correspond to the
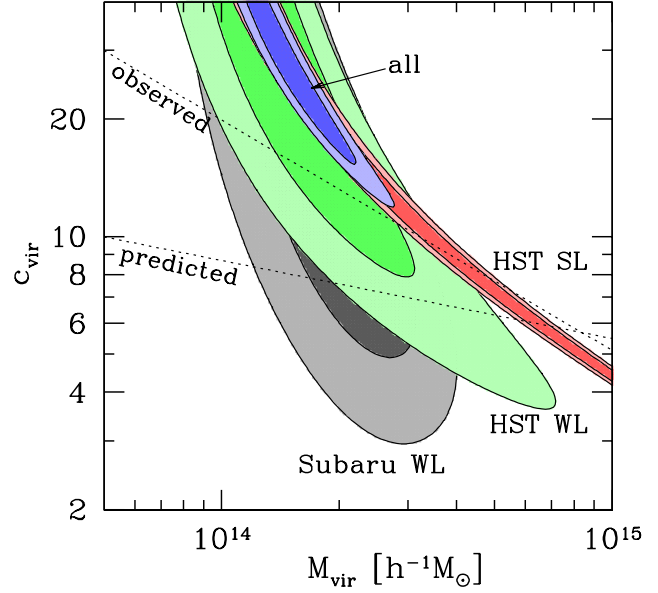
Figure 7. Constraints on the virial mass $M_{\rm vir}$ and the concentration parameter $c_{\rm vir}$ from the *HST* strong lensing (SL) analysis, *HST* weak lensing (WL), and Subaru weak lensing. Here the strong lensing constraint comes from the constraint on the Einstein radius, $\theta_E = 15''\!.2 \pm 0''\!.5$ for $z_s = 2.197$. The weak lensing constraints are derived from azimuthally averaged tangential shear profiles (see Figure 8). For each constraint, we show 1σ and 2σ ranges in this parameter space. The innermost contours show the combined constraints. Dotted lines indicate theoretically predicted and observed average mass-concentration relations for strong lensing clusters at $z \sim 0.5$ derived in Oguri et al. (2012).

Table 3. Best-fit mass and concentration parameters from combined lensing analysis (see also Figure 7). Errors correspond to 1σ confidence level.

constraints	$M_{\rm vir}$ $(10^{14} h^{-1} M_\odot)$	$c_{\rm vir}$
HST-WL	$1.35^{+0.89}_{-0.28}$	$30.9^{+8.9}_{-19.2}$
Subaru-WL	$2.00^{+0.70}_{-0.58}$	$11.5^{+14.2}_{-5.0}$
HST+Subaru-WL	$1.74^{+0.55}_{-0.45}$	$17.0^{+19.3}_{-6.5}$
HST-WL+SL	$1.32^{+0.63}_{-0.14}$	$34.7^{+5.1}_{-16.5}$
HST+Subaru-WL+SL	$1.55^{+0.40}_{-0.35}$	$25.7^{+14.1}_{-7.5}$

locations of G1a and G2. Thus, the two-dimensional weak lensing analysis indicates that the mass centroid is quite consistent with the average location of G1a and G2, but it is also consistent with the location of G1a or G2. In addition, we find that the shear map prefers an elongated mass distribution, with ellipticity $e = 0.72^{+0.13}_{-0.27}$ and position angle $\theta_e = 93^{+15}_{-11}$ deg. The position angle is in good agreement with both the position angle from strong lens modelling and that of the X-ray surface brightness profile.

5   COMBINED LENSING ANALYSIS

5.1   Radial density profile

We now combine the constraints from strong lensing (Section 3), *HST* weak lensing (Section 4), and Subaru weak lensing (Oguri et al. 2012) to derive constraints on the overall radial density profile of the lensing cluster. For the strong lensing result, we conservatively use the constraint on the Einstein radius of $\theta_E = 15''.2 \pm 0''.5$ for the source redshift of $z_s = 2.197$. For the *HST* and Subaru weak lensing data, we use the tangential shear profiles as constraints. We use these data to constrain the virial mass M_{vir} and the concentration parameter c_{vir}, assuming the NFW profile. The fit was performed for the mass range of $10^{13}h^{-1}M_\odot < M_{vir} < 10^{16}h^{-1}M_\odot$ and the concentration parameter range of $0.01 < c_{vir} < 39.8$.

We show constraints in the M_{vir}-c_{vir} plane in Figure 7, and summarize the resulting constraints on M_{vir} and c_{vir} in Table 3. The individual constraints are quite degenerate in this plane, but along slightly different directions so that we obtain a tighter constraint when combining the three constraints. We find $M_{vir} = 1.55^{+0.40}_{-0.35} \times 10^{14}h^{-1}M_\odot$ and $c_{vir} = 25.7^{+14.1}_{-7.5}$ from the combined analysis. The large concentration parameter value for the virial mass of $\sim 10^{14}h^{-1}M_\odot$ is in line with a recent study the M_{vir}-c_{vir} relation for strong lensing clusters (Oguri et al. 2012), which can be explained by the effect of baryon cooling and central galaxies (Fedeli 2012, but see also Duffy et al. 2010; McCarthy et al. 2010). However, the high concentration for this cluster should be interpreted cautiously given the complex core structure with two density peaks. We compare the tangential shear profiles with the best-fit model prediction in Figure 8. The three observational constraints are complementary and they are consistent with each other where they overlap. The best-fit NFW profile reproduces the observations for a wide range in radii.

Here we discuss several systematic effects that can affect our results. One is the redshift distribution of source galaxies for weak lensing analysis. Since the area of the field used for our weak lensing analysis is small, the redshift distribution in this field can be different from that of a mean distribution. Another effect is a multiplicative shear measurement bias, which is estimated to be $\lesssim 5\%$ for our Subaru weak lensing measurements. Also the *HST* weak lensing measurement probes the region $g \gtrsim 0.1$, where weak lensing measurements are less tested and therefore less reliable. To estimate a possible impact of these systematic errors, we consider an extreme situation where both the *HST* and Subaru weak lensing measurements are offset by $\pm 10\%$, and find the resulting shifts of the best-fit virial mass to be $\sim \pm 0.3 \times 10^{14}h^{-1}M_\odot$. The systematic error for this case is still comparable to the 1σ statistical error, implying that these systematic errors are not significant.

5.2   Comparison with X-ray mass

Ota et al. (2012) presented the *Chandra* X-ray analysis of SDSS J1029+2623, and derived a mass profile assuming hydrostatic equilibrium and isothermality. Figure 9 compares the X-ray mass profile from Ota et al. (2012) with the result of the combined lensing analysis in Section 5.1. We find
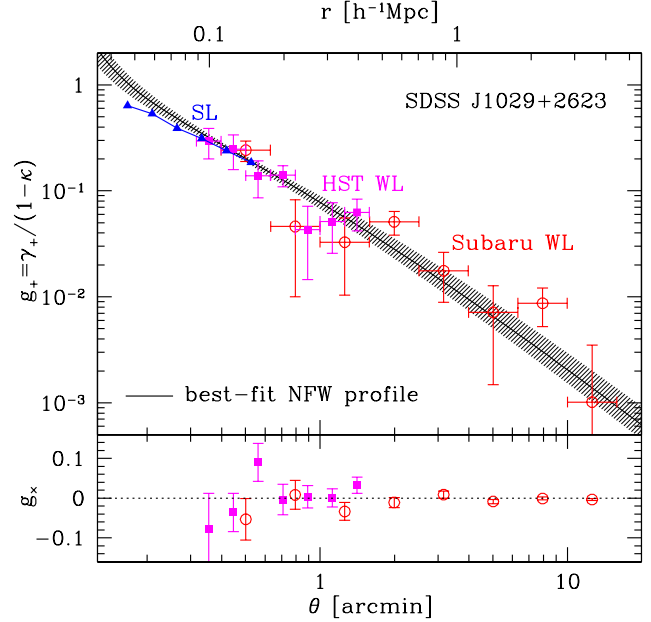

Figure 8. The tangential shear profile of the best-fit NFW model (*solid line*) as compared to the *HST* (*filled squares*) and Subaru (*open circles*) shear profiles and the reduced tangential shear g_+ predicted by the best-fit strong lens model (*filled triangles*) for $z_s = 1.1$. The best-fit shear profile is computed for $z_s = 1.1$, which is close to the effective source redshifts of $z_s = 1.06$ and 1.19 for the *HST* and Subaru weak lensing data. The shaded region indicates the 1σ range. The lower panel shows the radial profile of the $45°$ rotated component $g_\times$.

that the mass profiles derived from lensing and X-ray differ significantly with each other. While the enclosed masses agree at the radius of $\sim 100h^{-1}$kpc that roughly corresponds to the Einstein radius of this system, the enclosed masses inferred from the X-ray data are a factor of ~ 2 larger than those inferred from the combined lensing analysis at radii $r \gtrsim r_{2500}$, where r_{2500} is defined by the radius within which the average density is 2500 times the critical density at the cluster redshift. We note that the lensing derived mass implies an X-ray temperature and luminosity of $T \sim 2-3$ keV and $L_X \sim 10^{44}$ erg s^{-1} based on X-ray scaling relations (e.g., Dai, Kochanek, & Morgan 2007), which are significantly lower than the observed X-ray temperature of $T \sim 8.1$ keV and $L_X \sim 10^{45}$ erg s^{-1} (Ota et al. 2012). It is also worth noting that recent systematic weak lensing analysis of clusters at $z \lesssim 1$ found somewhat smaller weak lensing masses for a given X-ray temperature (Jee et al. 2011), which is qualitatively similar to our result, although the difference appears to be much smaller than found here.

There are several effects that can lead to a systematic difference between the X-ray and lensing masses. One of the most significant effects for X-ray cluster mass measurements is any violation of the assumption of hydrostatic equilibrium due the presence of non-thermal pressure support. However, this effect typically leads to an underestimate of the X-ray mass, particularly in the outskirts of clusters (e.g., Mahdavi et al. 2008; Zhang et al. 2010), and would only make the discrepancy larger. Another possibility is our assumption of isothermality. In fact, X-ray tempera-
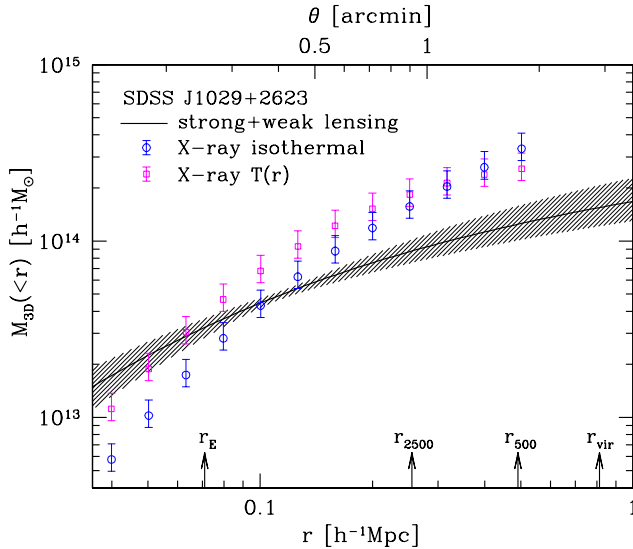
Figure 9. The enclosed mass within a sphere of radius r from the *Chandra* X-ray analysis and from the combined lensing analysis (*solid line*; see Section 5.1), assuming a spherically symmetric mass distribution. For the *Chandra* X-ray analysis, we show the results for the isothermal β-model from Ota et al. (2012, *open circle*) and a β-model with the temperature profile of Burns et al. (2010, *open squares*). The errors in the X-ray masses come from the statistical errors in the X-ray temperature measurement. For reference, we also show the cluster radii $r_{\rm vir}$, r_{500}, and r_{2500}, which are computed from the best-fit lensing mass profile, as well as the Einstein radius $r_{\rm E}$.

tures generally decrease at large radii, but our *Chandra* data were not sensitive enough to detect such a decrease. To examine this effect, we adopt a temperature profile obtained from hydrodynamic simulations (Burns et al. 2010), which appears to be consistent with recent X-ray observations of cluster outskirts (e.g., Akamatsu et al. 2011), and recalculate the X-ray mass profile. While this reduces the difference at the center, it cannot explain the overall difference between the X-ray and lensing masses. It is also difficult to explain the difference by a triaxial halo, because the large $c_{\rm vir}$ implies that the major-axis of the cluster is more likely to be aligned with the line-of-sight direction, in which case the lensing mass should be overestimated (Oguri et al. 2005).

The most likely explanation for the mass discrepancy is shock heating of the intracluster gas during a merger. Numerical simulations show significant boosts of X-ray luminosity and temperature ~ 1 Gyr after mergers, which can lead to significant overestimates of X-ray masses (e.g., Ricker & Sarazin 2001; Takizawa, Nagino, & Matsushita 2010; Rasia et al. 2011; Nelson et al. 2012). Observationally, there are several clusters showing signs of ongoing mergers that also have significantly higher X-ray masses than lensing masses (e.g., Okabe & Umetsu 2008; Okabe et al. 2011; Soucail 2012). Indeed the lensing cluster of SDSS J0129+2623 shows hints of an ongoing merger, including the bimodal nature of cluster cores, the complex X-ray morphology, a possible offset between mass and X-ray centroids, and the large G1-G2 velocity difference. In addition, a line-of-sight merger can naturally explain the high concentration estimates for this cluster (e.g., King & Corless

2007). Spectroscopy of many more cluster member galaxies is needed to understand this complex cluster further. If this interpretation is correct, the agreement between X-ray and lensing masses near the Einstein radius might just be a coincidence, in that both the masses are overestimated by merger shock heating and halo elongation along the line-of-sight, respectively.

5.3 Gas-to-mass ratio

Another useful cross-check of our interpretation is provided by the gas-to-mass ratio, $f_{\rm gas}(< r) = M_{\rm gas}(< r)/M_{\rm tot}(< r)$, because it is expected to roughly match the cosmic baryon fraction $\Omega_b/\Omega_M \simeq 0.167$ (Komatsu et al. 2011) for massive clusters, and also because $M_{\rm gas}$ appears to be the most promising cluster mass proxy (Okabe et al. 2010; Fabjan et al. 2011). We use the gas mass profile implied by the isothermal β-model from Ota et al. (2012) to estimate a gas mass within the radius $r_{500} = 0.49h^{-1}$Mpc from the combined lensing analysis. We find a gas mass fraction of $f_{\rm gas}(< r_{500}) = 0.060 \pm 0.014$, which is significantly smaller than the cosmic baryon fraction. The gas mass fraction is slightly increased to $f_{\rm gas}(< r_{500}) = 0.077 \pm 0.017$ when the temperature profile of Burns et al. (2010) is used to derive the total mass, but the value is still low. On the other hand, when the lensing mass estimate is used for the total mass, we obtain a gas mass fraction of $f_{\rm gas}(< r_{500}) = 0.158 \pm 0.038$, which is in better agreement with $f_{\rm gas}$ estimates in other clusters (e.g., Vikhlinin et al. 2006; Allen et al. 2008; Ettori et al. 2009; Umetsu et al. 2009; Dai et al. 2010b; Okabe et al. 2010; Pratt et al. 2010). This further supports our interpretation that the X-ray derived mass is significantly overestimated due to shock heating of the intracluster gas during a merger.

6 IMPLICATIONS FOR THE FLUX RATIO ANOMALY

6.1 Searching for substructures

Our strong lens models in Section 3 confirm the earlier claims (Oguri et al. 2008; Kratzer et al. 2011; Ota et al. 2012) that a substructure is required to explain the anomalous flux ratios between the three quasar images. In order to search for possible substructures, we subtract the quasars from each *HST* image and carefully examine the residual images. In the ACS/F814W image, we find a probable galaxy in the vicinity of image C, as shown in Figure 10. We fit the galaxy, named GX, assuming a Sersic profile with index $n = 4$ using *galfit* (Peng et al. 2002), and find the centroid of the galaxy to be $\sim 0\rlap{.}''4$ from image C (R.A.=157.309482, Decl.=26.394158), and that the magnitude of the galaxy is $m_{F814W} \sim 24.3$. These estimates have large uncertainties due to the significant residuals from the subtraction of image C (see Figure 10). In the strong lens models we estimate that galaxy G2 with $m_{F814W} = 19.62$ has a velocity dispersion of $\sigma = 291$ km s^{-1}, which implies a velocity dispersion of $\sigma \simeq 100$ km s^{-1} for GX assuming $\sigma \propto L^{1/4}$. The inferred velocity dispersion of GX is large enough to mean that it will significantly perturb the flux of image C.
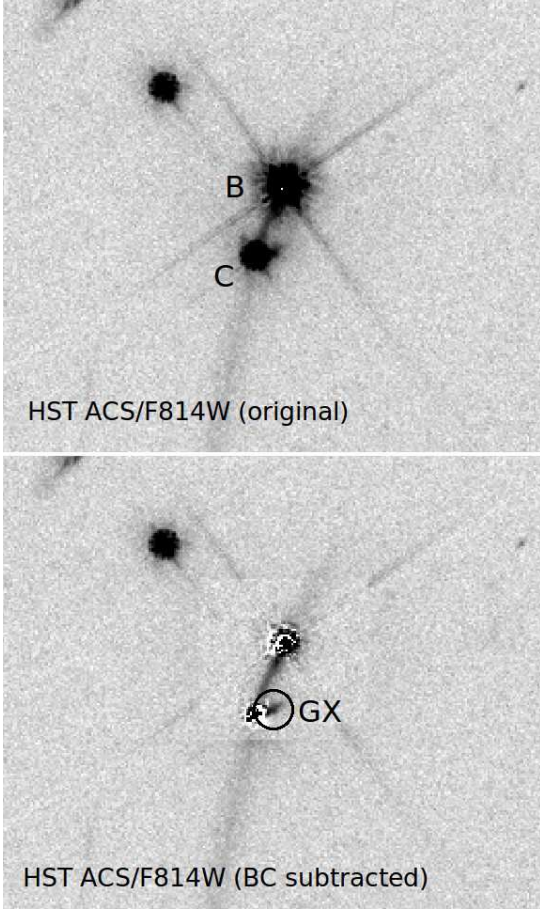
Figure 10. *Top:* The *HST* ACS/F814W image around the merging quasar images B and C. *Bottom:* The image after subtracting quasar images. The probable galaxy GX is marked by a circle. We used a nearby star as a PSF template.

6.2 Mass modelling

We next explore the impact of GX on the strong lens models. Here we assume the same mass model as used in Section 3, but include an additional component modelled by pseudo-Jaffe model at the location of GX with the velocity dispersion and the cutoff radius as free parameters. We find that the best-fit model has $\chi^2 = 7.5$ including the flux ratios, which is significantly smaller than the χ^2 of the original best-fit model ($\chi^2 = 30$) and the large improvement is due to improvements in fitting the flux ratios. This confirms the argument by Kratzer et al. (2011) that a substructure in the vicinity of image C can explain the anomalous flux ratio. The velocity dispersion of galaxy GX in the best-fit model is $\sigma = 76\,\mathrm{km\,s^{-1}}$, which is broadly consistent with the velocity dispersion expected from the magnitude. We summarize the flux ratios of the best-fit mass models, as well as the observed flux ratios at various wavelengths, in Figure 11.

The best-fit models with and without galaxy GX have quite different total magnification factors. For the model without galaxy GX, the magnifications for the quasar images are $\mu_A = 6.8$, $\mu_B = 82.9$, and $\mu_C = 94.7$, whereas for the model with galaxy GX, they are $\mu_A = 10.4$, $\mu_B = 13.8$, and $\mu_C = 5.5$. The total magnifications are $\mu_{\rm tot} = 184.4$ for the model without GX and $\mu_{\rm tot} = 29.7$ for the model with GX.
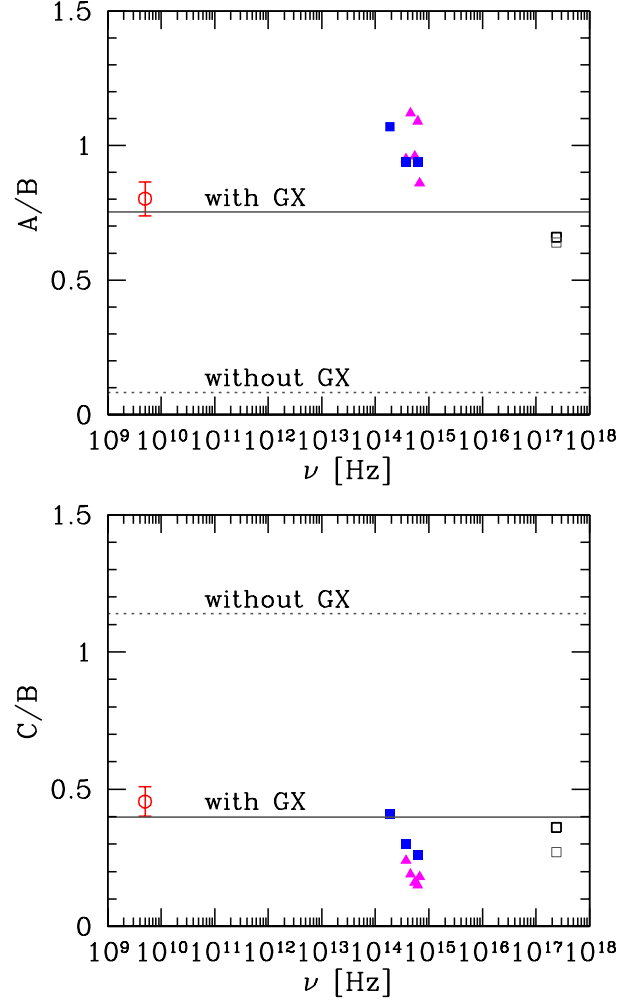

Figure 11. Flux ratios of the best-fit strong lens models with (*solid*) and without (*dotted*) galaxy GX are compared with the observed flux ratios in different wavelengths. The top panel shows the flux ratio between images A and B, while the bottom panel shows the flux ratio between images B and C. Observed flux ratios are obtained from the VLA radio image (*open circle*) by Kratzer et al. (2011), the *HST* images (*filled squares*) presented in this paper, ground-based optical images (*filled triangles*) by Oguri et al. (2008), and the *Chandra* X-ray image with (*black open square*) and without (*grey thin open square*) the absorption correction by Ota et al. (2012).

Substructures near merging image pairs can have large effects on time delays between the images (Oguri 2007; Keeton & Moustakas 2009). This was indeed the case for SDSS J1004+4112 (Inada et al. 2003), another large-separation lensed quasar produced by a massive cluster of galaxies, in which a time delay between the smallest image pair was found to be affected by the perturbation from nearby cluster member galaxies (Fohlmeister et al. 2007). We find that our best-fit mass models predict the time delay between image B and C as $\Delta t_{BC} = 0.7$ day for the model without GX, and $\Delta t_{BC} = 7.7$ day for the model with GX. In both cases, image B leads image C as expected from the image geometry. On the other hand, GX has a negligible impact on the AB time delay. The large difference of Δt_{BC}

implies that measuring the BC time delay will greatly help constrain the substructure that affects these quasar images.

6.3 Morphology of the lensed host galaxy

An interesting feature of SDSS J1029+2623 is the presence of a highly elongated host galaxy, which is most prominent in the *HST* WFC3/F160W image. Here we examine whether our best-fit models reproduce the morphology of the lensed host galaxy. To do so, we include the host galaxy modelled by a Sersic profile, and optimise the parameters (the total flux, size, ellipticity and position angle) of the Sersic profile assuming the best-fit mass models, both with and without galaxy GX. We fix the centroid of the host galaxy in the source plane to the best-fit source quasar position. The optimisation is conducted on a pixel-by-pixel basis using a binned *HST* WFC3/F160W image with a pixel scale of $0\farcs2$. The model-predicted images are convolved with a PSF assuming a Moffat profile with a FWHM of $0\farcs18$ and $\beta = 2$, before comparing with the observed image.

The result is shown in Figure 12. We find that our best-fit models reproduce the morphology of the lensed host galaxy quite well. The match is better for the model with galaxy GX, which further supports its presence and the significant impact of galaxy GX on this lens system. In particular, we note that the observed host galaxy shows signs of a kink near image C (see also Figure 10) which is reproduced in our best-fit model with galaxy GX. We estimate the total magnitude of the lensed host galaxy as $m_{F160W} \sim 18.5$ and the lensing-corrected original magnitude as $m_{F160W} \sim 22.4$. The best-fit effective radius and Sersic index are $r_e = 0\farcs27$ and $n = 0.88$, respectively. The presence of galaxy GX has a negligible impact on the total magnification of the host galaxy. The highly magnified host galaxy provides a very unique opportunity to study the relation between supermassive black holes and host galaxies in a distant quasar (e.g., Peng et al. 2006; Ross et al. 2009).

7 CONCLUSION

We have presented a detailed gravitational lensing analysis of SDSS J1029+2623, the largest-separation lensed quasar system known (image separation $\theta = 22\farcs5$, lens redshift $z_l = 0.584$, and source redshift $z_s = 2.197$), based on new *HST* ACS and WFC3 observations. With the help of the high-resolution *HST* images, we have identified 6 new multiply imaged galaxies in addition to the three-image lensed quasar. Based on the newly identified multiple images, we construct an accurate mass model at the cluster core. We find dark haloes with roughly similar masses should be associated with both of the bright central galaxies. The strong lens models allow us to estimate the Einstein radius of the overall lens as $\theta_E = 15\farcs2 \pm 0\farcs5$. Our best-fit strong lens model successfully reproduces the observed time delay between quasar image A and B which was recently measured by Fohlmeister et al. (2012). We have also detected a significant weak lensing signal in the *HST* ACS image. The mass map obtained from the strong and weak lensing analysis appears to be different from the X-ray surface brightness profile (Ota et al. 2012), suggesting an ongoing merger.

We have constrained the radial mass profile of the cluster from the *HST* strong and weak lensing analysis as well as the Subaru weak lensing result (Oguri et al. 2012). The combination of the three lensing constraints enabled a robust determination of the radial mass profile, with a best-fit virial mass of $M_{vir} = 1.55^{+0.40}_{-0.35} \times 10^{14} h^{-1} M_\odot$ and a concentration parameter of $c_{vir} = 25.7^{+14.1}_{-7.5}$. The lensing derived mass profile differs significantly from the mass profile inferred from the *Chandra* X-ray data assuming hydrostatic equilibrium and isothermality, such that the X-ray mass is a factor of ~ 2 larger at outer radii of $r_{2500} \lesssim r \lesssim r_{500}$. We ascribe the mass discrepancy to shock heating of the intracluster gas during a merger, which leads to significant overestimates of X-ray cluster masses. Our interpretation is supported by the small cluster gas mass fraction implied by the X-ray estimates of total masses, with $f_{gas}(< r_{500}) = 0.060 \pm 0.014$, as compared to $f_{gas}(< r_{500}) = 0.158 \pm 0.038$ using the lensing estimate of the total mass. A merger along the line-of-sight can also explain the high concentration parameter value found for the cluster, and is consistent with the large velocity difference of $\Delta v \simeq 2800$ km s^{-1} between the bright central galaxies G1 and G2 (Oguri et al. 2008). The merger interpretation for this strong lens selected cluster, in turn, may imply a large impact of cluster mergers on strong lensing cross sections (e.g., Torri et al. 2004; Fedeli et al. 2006; Redlich et al. 2012).

If the merger interpretation is correct, one question is whether the large velocity difference is compatible with the theoretical expectations. There are several works examining the distribution of velocity differences between close halo pairs (e.g., Hayashi & White 2006; Lee & Komatsu 2010; Thompson & Nagamine 2012), which indicate that a relative velocity of $\Delta v \simeq 2800$ km s^{-1} for haloes with $M_{vir} \gtrsim 10^{14} h^{-1} M_\odot$ is quite rare. Therefore, it is important to make more quantitative comparisons taking account of the redshift and mass of SDSS J1029+2623 in order to check whether this large velocity difference can be explained in the framework of the standard Λ-dominated cold dark matter cosmology.

We confirm the earlier claims (Oguri et al. 2008; Kratzer et al. 2011; Ota et al. 2012) that the observed flux ratios between the quasar images cannot be reproduced by strong lens models without any substructure in the vicinity of the quasar images. From a careful inspection of the *HST* images, we have identified a probable galaxy GX which is only $\sim 0\farcs4$ from quasar image C. A mass model including GX successfully reproduces the observed flux ratios, thereby resolving the flux ratio anomaly. While GX has little effect on the AB time delay measured by Fohlmeister et al. (2012), the BC time delay is a sensitive probe of its effects. We have also checked the morphology of the lensed host galaxy and found a very good agreement with the model prediction. The best-fit model predict the total magnification of the quasar as $\mu_{tot} = 30$ and that of the lensed quasar host galaxy as $\mu_{tot} = 35$.

Finally we note that additional observations are essential to better characterise this unique quasar-cluster lens system. For instance, the measurement of the time delay between quasar images B and C will be important for constraining the properties of galaxy GX. The merger interpretation for the lensing cluster should be checked by extensive spectroscopy of cluster member galaxies. Spectroscopy of the
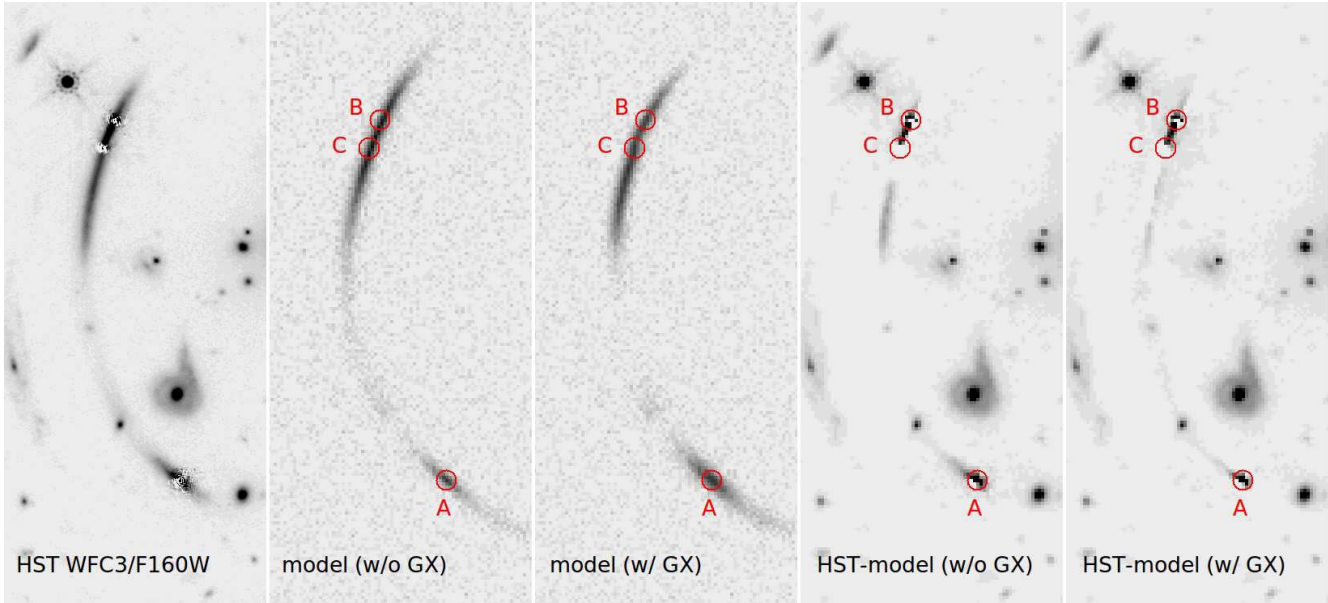
Figure 12. *First panel:* The *HST* WFC3/F160W image showing the lensed host galaxy of the quasar at $z_s = 2.197$. All the quasar components are subtracted using a nearby star as a PSF template. *Second panel:* The morphology of the lensed host galaxy in the best-fit mass model without galaxy GX from Section 3. The locations of the three quasar images are marked by circles. *Third panel:* Similar to the middle panel, but the case for the best-fit model with galaxy GX. *Fourth panel:* The *HST* WFC3/F160W image after subtracting the best-fit host galaxy model without galaxy GX. *Fifth panel:* The *HST* WFC3/F160W image after subtracting the best-fit host galaxy model with galaxy GX. Note that all the panels are presented with the same grey-scale stretch.

multiply imaged galaxies identified in this paper is important both for more robust mass models and for possible cosmological constraints (e.g., Jullo et al. 2010). Furthermore, the highly elongated host galaxy makes this lens system an ideal site to conduct detailed explorations of a quasar host galaxy at $z \sim 2$ (e.g., Ross et al. 2009).

REFERENCES

Akamatsu H., Hoshino A., Ishisaki Y., Ohashi T., Sato K., Takei Y., Ota N., 2011, PASJ, 63, 1019
Allen S. W., Rapetti D. A., Schmidt R. W., Ebeling H., Morris R. G., Fabian A. C., 2008, MNRAS, 383, 879
Bar-Kana R., 1996, ApJ, 468, 17
Bertin E., Arnouts S., 1996, A&AS, 117, 393
Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, 1503
Broadhurst T., Takada M., Umetsu K., Kong X., Arimoto N., Chiba M., Futamase T., 2005, ApJ, 619, L143
Burns J. O., Skillman S. W., O'Shea B. W., 2010, ApJ, 721, 1105
Coe D., Benítez N., Sánchez S. F., Jee M., Bouwens R., Ford H., 2006, AJ, 132, 926
Cohn J. D., Kochanek C. S., McLeod B. A., Keeton C. R., 2001, ApJ, 554, 1216
Dai X., Kochanek C. S., Morgan N. D., 2007, ApJ, 658, 917
Dai X., Kochanek C. S., Chartas G., Kozłowski S., Morgan C. W., Garmire G., Agol E., 2010a, ApJ, 709, 278
Dai X., Bregman J. N., Kochanek C. S., Rasia E., 2010b, ApJ, 719, 119
Duffy A. R., et al., 2010, MNRAS, 405, 2161
Ettori S., Morandi A., Tozzi P., Balestra I., Borgani S., Rosati P., Lovisari L., Terenziani F., 2009, A&A, 501, 61
Fabjan D., Borgani S., Rasia E., Bonafede A., Dolag K., Murante G., Tornatore L., 2011, MNRAS, 416, 801
Fedeli C., Meneghetti M., Bartelmann M., Dolag K., Moscardini L., 2006, A&A, 447, 419
Fedeli C., 2012, MNRAS, 424, 1244
Fohlmeister J., et al., 2007, ApJ, 662, 62

Fohlmeister J., Kochanek C. S., Falco E. E., Wambsganss J., Oguri M., Dai X., 2012, arXiv:1207.5776

Hayashi E., White S. D. M., 2006, MNRAS, 370, L38

Inada N., et al., 2003, Natur, 426, 810

Inada N., et al., 2006, ApJ, 653, L97

Jee M. J., et al., 2011, ApJ, 737, 59

Jullo E., Kneib J.-P., Limousin M., Elíasdóttir Á., Marshall P. J., Verdugo T., 2007, NJPh, 9, 447

Jullo E., Natarajan P., Kneib J.-P., D'Aloisio A., Limousin M., Richard J., Schimd C., 2010, Sci, 329, 924

Kaiser N., Squires G., 1993, ApJ, 404, 441

Kaiser N., Squires G., Broadhurst T., 1995, ApJ, 449, 460

Keeton C. R., Moustakas L. A., 2009, ApJ, 699, 1720

King L., Corless V., 2007, MNRAS, 374, L37

Koekemoer A. M., Fruchter A. S., Hook R. N., Hack W., 2002, hstc.conf, 337

Komatsu E., et al., 2011, ApJS, 192, 18

Kratzer R. M., Richards G. T., Goldberg D. M., Oguri M., Kochanek C. S., Hodge J. A., Becker R. H., Inada N., 2011, ApJ, 728, L18

Lee J., Komatsu E., 2010, ApJ, 718, 60

Li G. L., Mao S., Jing Y. P., Lin W. P., Oguri M., 2007, MNRAS, 378, 469

Mahdavi A., Hoekstra H., Babul A., Henry J. P., 2008, MNRAS, 384, 1567

McCarthy I. G., et al., 2010, MNRAS, 406, 822

Medezinski E., et al., 2007, ApJ, 663, 717

Medezinski E., Broadhurst T., Umetsu K., Oguri M., Rephaeli Y., Benítez N., 2010, MNRAS, 405, 257

Meneghetti M., Argazzi R., Pace F., Moscardini L., Dolag K., Bartelmann M., Li G., Oguri M., 2007, A&A, 461, 25

Minor Q. E., Kaplinghat M., 2008, MNRAS, 391, 653

Miyazaki S., et al., 2002, PASJ, 54, 833

Motta V., Mediavilla E., Falco E., Munoz J. A., 2012, ApJ, 755, 82

Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493

Nelson K., Rudd D. H., Shaw L., Nagai D., 2012, ApJ, 751, 121

Oguri M., 2007, ApJ, 660, 1

Oguri M., 2010, PASJ, 62, 1017

Oguri M., Keeton C. R., 2004, ApJ, 610, 663

Oguri M., Takada M., Okabe N., Smith G. P., 2010, MNRAS, 405, 2215

Oguri M., Takada M., Umetsu K., Broadhurst T., 2005, ApJ, 632, 841

Oguri M., Bayliss M. B., Dahle H., Sharon K., Gladders M. D., Natarajan P., Hennawi J. F., Koester B. P., 2012, MNRAS, 420, 3213

Oguri M., et al., 2008, ApJ, 676, L1

Okabe N., Umetsu K., 2008, PASJ, 60, 345

Okabe N., Zhang Y.-Y., Finoguenov A., Takada M., Smith G. P., Umetsu K., Futamase T., 2010, ApJ, 721, 875

Okabe N., Bourdin H., Mazzotta P., Maurogordato S., 2011, ApJ, 741, 116

Ota N., et al., 2012, ApJ, 758, 26

Peng C. Y., Ho L. C., Impey C. D., Rix H.-W., 2002, AJ, 124, 266

Peng C. Y., Impey C. D., Rix H.-W., Kochanek C. S., Keeton C. R., Falco E. E., Lehár J., McLeod B. A., 2006, ApJ, 649, 616

Pratt G. W., et al., 2010, A&A, 511, A85

Rasia E., Mazzotta P., Evrard A., Markevitch M., Dolag K., Meneghetti M., 2011, ApJ, 729, 45

Redlich M., Bartelmann M., Waizmann J.-C., Fedeli C., 2012, A&A, in press (arXiv:1205.6906)

Ricker P. M., Sarazin C. L., 2001, ApJ, 561, 621

Ross N. R., Assef R. J., Kochanek C. S., Falco E., Poindexter S. D., 2009, ApJ, 702, 472

Schrabback T., et al., 2010, A&A, 516, A63

Soucail G., 2012, A&A, 540, A61

Takizawa M., Nagino R., Matsushita K., 2010, PASJ, 62, 951

Thompson R., Nagamine K., 2012, MNRAS, 419, 3560

Torri E., Meneghetti M., Bartelmann M., Moscardini L., Rasia E., Tormen G., 2004, MNRAS, 349, 476

Umetsu K., Broadhurst T., 2008, ApJ, 684, 177

Umetsu K., et al., 2009, ApJ, 694, 1643

Vikhlinin A., Kravtsov A., Forman W., Jones C., Markevitch M., Murray S. S., Van Speybroeck L., 2006, ApJ, 640, 691

Zhang Y.-Y., et al., 2010, ApJ, 711, 1033